Claims

What is claimed is:

1. A method of providing speaker adaptation in speech recognition, said method comprising the steps of:

providing at least one speech recognition model;

accepting speaker data;

generating a word lattice based on the speaker data; and

adapting at least one of the speaker data and the at least one speech recognition model in a manner to maximize the likelihood of the speaker data with respect to the generated word lattice.

2. The method according to Claim 1, wherein said step of generating a word lattice comprises generating a maximum a-posteriori probability word lattice.

3. The method according to Claim 2, wherein said step of generating a maximum a-posteriori probability word lattice comprises:

determining posterior state occupancy probabilities for each state in the speaker data at each time;

determining posterior word occupancy probabilities by summing over all states interior to each word in the speaker data; and

determining at least one likeliest word at each frame of the speaker data.

4. The method according to Claim 2, wherein said step of generating a word lattice further comprises connecting word traces into a lattice.

5. The method according to Claim 1, further comprising the step of discarding interpretations associated with low confidence.

6. The method according to Claim 5, wherein said discarding step comprises determining posterior phone probability.

7. The method according to Claim 1, wherein said adapting step comprises performing maximum likelihood linear regression on the speaker data.

8. An apparatus for providing speaker adaptation in speech recognition, said apparatus comprising:

at least one speech recognition model;

an accepting arrangement which accepts speaker data;

a lattice generator which generates a word lattice based on the speaker data; and

a processing arrangement which adapts at least one of the speaker data and the at least one speech recognition model in a manner to maximize the likelihood of the speaker data with respect to the generated word lattice.

9. The apparatus according to Claim 8, wherein said generator is adapted to generate a maximum a-posteriori probability word lattice.

10. The apparatus according to Claim 9, wherein said generator is adapted to:

determine posterior state occupancy probabilities for each state in the speaker data at each time;

determine posterior word occupancy probabilities by summing over all states interior to each word in the speaker data; and

determine at least one likeliest word at each frame of the speaker data.

11. The apparatus according to Claim 9, wherein said generator is further adapted to connect word traces into a lattice.

12. The apparatus according to Claim 8, further comprising a discarding arrangement which discards interpretations associated with low confidence.

13. The apparatus according to Claim 12, wherein said discarding arrangement is adapted to determine posterior phone probability.

14. The apparatus according to Claim 8, wherein said processing arrangement is adapted to perform maximum likelihood linear regression on the speaker data.

15. A program storage device readable by machine, tangibly embodying a program of instructions executable by the machine to perform method steps for providing speaker adaptation in speech recognition, said method comprising the steps of:

providing at least one speech recognition model;

accepting speaker data;

generating a word lattice based on the speaker data; and

adapting at least one of the speaker data and the at least one speech recognition model in a manner to maximize the likelihood of the speaker data with respect to the generated word lattice.